AMERICAN

STATISTICAL ASSOCIATION.

NEW SERIES, No. 58. JUNE, 1902.

WHAT IS THE CENTER OF AN AREA, OR THE CENTER OF A POPULATION?

By John F. Hayford.

There are many curious misconceptions and absurdities in common circulation in regard to the meaning of the expressions "center of area" and "center of population." The educated, as well as the relatively ignorant, man seems to be puzzled in regard to these terms. No accurate definitions of them exist in current usage, nor can satisfactorily accurate definitions be found in the dictionaries. This paper is an attempt to help in the bringing about of a clear conception of these terms and to remove some of the real and apparent absurdities which arise when they are used.

The center of an area is a point such that if a perpendicular be dropped from each point to *any* one line drawn through said center the sum of the perpendiculars lying on one side of the line will be equal to the sum of the perpendiculars on the other side. The point fixed upon by this definition is the same that would be obtained by representing the area by a thin plate of metal of uniform thickness and density and taking the center of gravity, or point about which the plate will balance, to represent the center of area.

The mathematical process of locating a center of area is not difficult. A line is first drawn across the area in any

assumed direction. For convenience it is well to draw this line in such a position that it will pass as nearly through the center as that point can be located by mere inspection. The area is then divided up into small areas each of which lies completely on one side or the other of the selected line. The sum of the products is then taken separately for each side of the line, of each area into distance of the center from the line. The difference of these two sums of products divided by the total area is then the distance which the selected line must be shifted sidewise to make it pass through the center required, and the direction of the shift must be toward the side on which the sum of the products was the greater. In the new position of the line the sums of the products will be the same on the two sides of the line. A second trial line is now drawn in some other direction than the first, and preferably at right angles to it, and the whole process repeated. Two lines will thus have been obtained each of which pass through the required center which therefore must be at their intersection.

It may be noted by the careful reader that in the above process it is assumed that the center of each of the small component areas into which the total area is divided may be determined. If the component areas are regular figures, such as squares or rectangles, there is no difficulty in locating the center of area of each. If they are not regular figures the position of the center of each may be estimated by eye and the accuracy of the final result tested by some indirect process, or these centers may each be located by a separate process similar to that which is used for the whole area. Obviously the smaller each component area is the more accurately can its center be located by eye, and the less is the effect upon the final result of a given error in the location of each such secondary center. Hence any desired degree of accuracy may be obtained by sufficiently minute subdivision.

There are four properties or characteristics of the center of area defined and located, as indicated above, which are important and interesting.

First, as already stated, it corresponds to the center of gravity if the area be considered to be represented by a thin plate of uniform thickness and density.

Second, The point as thus defined is always unique, that is, there is but one center under the definition for each area. After the center has been located by finding two lines each fulfilling the condition that the sum of the perpendiculars on the two sides shall be the same, it will be found that for any other line through this center the condition is also fulfilled. Or, conversely, the center obtained for any given area is absolutely independent of the directions of the two lines used to fix the center. To illustrate, if in a given county there is a man upon the center of each square mile of area and no other population, and if every man travels to the center of area of the county by first going directly east or west to the meridian passing through the center and then following the meridian to the center, the total distance traveled to the westward will be equal to the total eastward travel, and the total distance traveled to the northward will be equal to the total to the southward. The totals will also be equal in pairs if the traveling be done parallel to any other pair of rectangular axes, let us say, northwest and southeast and northeast and southwest.

Third, The sum of the squares of the perpendiculars dropped from each point of the area to a line passing through the center is smaller than for any other line parallel to it.

Fourth, The sum of the squares of the lines drawn from each point of the area to the center is smaller than the sum of the squares of such lines drawn to any other point. Or, in other words, if it be assumed as before that there is a man on the center of each square mile of a county and each man goes to the center over a straight course the sum

of the squares of the distances traveled will be less than if they so traveled to any other point than the center.

The third and fourth properties of the center are chiefly interesting from a mathematical point of view because they indicate a close analogy between the center so defined and the mean of number of quantities.

Aside from the definition of center given above the most common definition set forth is what may be called the median or bisector definition. The center is defined as being the point at which two lines which bisect the area intersect. According to this definition exactly one half of a county lies on each side of a meridian drawn through its center and exactly one half on each side of an east and west line drawn through its center.

Many people believe the center as thus defined to be the same as that given by the definition advocated in this paper and which for convenience we will call the gravity definition as contrasted with the bisector definition. The two definitions do not, however, give the same point except in a few cases. In some cases the two points are far apart.

The bisector definition is often put forward by people who know that it gives a point different from the gravity definition, and who put forward the contention that the gravity definition is absurd. They say, and it is true, that under the gravity definition the addition or subtraction of a given area far from the center moves the center much more than if such area were added or subtracted near the center. They say it is absurd that distant areas should have so much more effect upon the location of the center than near areas. This is not really an absurdity. The quickest way, however, of convincing a man who advocates the bisector definition because of this alleged absurdity in the gravity definition is not by attempting to show him that the alleged absurdity he has pointed out is merely apparent and not real, but by showing him the real absurdities in the bisector definition.

The absurdities of the bisector definition may quickly be brought out by such a diagram as Figure 1.

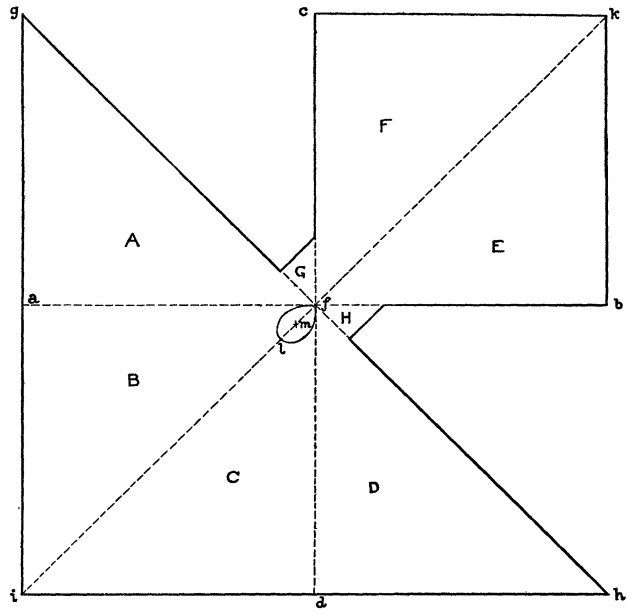

FIG. 1.

The area shown is a square with two portions cut away. Of the triangles into which the area is divided by dotted lines the large triangles *A*, *B*, *C*, *D*, *E* and *F* are all equal and the small triangles *G* and *H* are also equal. The lines *ab* and *cd* each bisect the area and, therefore, by the bisector definition the point *f* is the center. Look now at the line *gh*, which passes through the alleged center *f*. *Four* large triangles, *A*, *B*, *C*, and *D*, lie on one side of *gh* while on the other side there are only *two* large triangles, *E* and *F*, and the two small triangles, *G* and *H*. This line *gh*, which passes through the alleged center under the bisector definition, has nearly two-thirds of the whole area on one side of it. Obviously there is an absurdity here.

If the line *gh* be moved parallel to itself until it does bisect the area it will be found to pass through the point *l*. The line *ik* bisects the area. Hence under the bisector definition *l* is the center of area as well as *f*. In other words, this one area according to the bisector definition has two centers, which is an obvious absurdity.

These two centers under the bisector definition have been obtained by using two pairs of rectangular bisectors of which the second pair are at an angle of forty-five degrees with the first pair. It can be shown that for this area a different center will be obtained for each pair of bisectors used. If only bisectors which are at right angles are used it will be found that this area has under the bisector definition an infinite number of centers all lying in the periphery of an oval having its extreme points at *f* and *l*.

If the advocate of the bisector definition is not sufficiently convinced of its absurdity by the fact that it gives each irregular area an infinite number of centers instead of one he may be invited to consider the matter from another point of view.

It is axiomatic that if any portion of an area be moved the center should also move provided the movement of one portion of the area was not accompanied by a simultaneous, symetrical movement of a similar portion of the area.

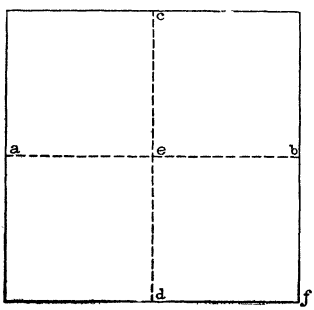

FIG. 2.

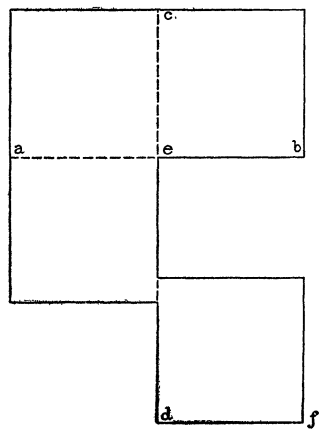

FIG. 3.

Figure 2 represents a square of which the center under the bisector definition is located at *e*, the intersection of the bisectors *ab* and *cd*. Figure 3 represents the same area after the portion *debf* has been moved a considerable distance. The lines *ab* and *cd* are, however, still bisectors and under the bisector definition the center is still at the point *e*. In fact the portion *debf* might have been moved about and distorted to any extent whatever provided no portion of it were carried across either bisector, *cd* or *ab*, without moving the center according to the bisector definition.

Both the absurdities pointed out in connection with the bisector definition are real and are so contradictory to the most elementary and fundamental conceptions of the word center that further comment is unnecessary, except perhaps to call attention by way of contrast to the fact that under the gravity definition each area has but one center, and any movement of any portion of the area will move the center unless it is accompanied by some symetrical movement of another portion of the area.

In Figure 1 the center according to the gravity definition lies within the oval at the point marked *m*.

What is the center of population of the United States? It is defined correctly as follows in Bulletin No. 62 of the Twelfth Census: "The center of population is the center of gravity of the population of the country, each individual being assumed to have the same weight." This corresponds exactly to the gravity definition of a center of area.

The process of locating the center of population as outlined in this bulletin also corresponds exactly to the process of locating the center of area given in the first part of this paper.

Various objectors say that this definition of the center of population is absurd. The greatest apparent absurdity urged in objection is that a man at a distance from the center has much more influence upon the center than a man near at hand. The death of a man in San Francisco moves the

center of population more than the death of a hundred men in Indiana living within ten miles of the center. These objectors advocate the use of the bisector definition for determining the center of population.

In Bulletin No. 62 a concession to these objectors has been made and a "median point" of the population has been located. The "median point" is defined as "the point of intersection of the line dividing the population equally north and south with the line dividing it equally east and west" This corresponds exactly to the bisector definition of the center of an area with the exception that this defintion has carefully been narrowed by limiting the direction of the bisectors to north and south and east and west so that it will give but one median point. If this restriction be removed so that the two bisectors may be taken in any direction it will be found that the population of the United States has an infinite number of "median points" just as the area in Figure 1 has an infinite number of centers according to the bisector definition.

It may be urged that the center of population according to the gravity definition and all the possible median points are all confined to so small an area that it is immaterial which is taken as the center. This proposition will not stand examination.

The median point for 1900 as defined in Bulletin No. 62 is eighty miles to north-eastward of the center according to the gravity definition. This distance is greater than the total movement of the latter point in the last two decades. This distance between the center and the median point of a population depends, firstly, upon the shape of the area over which the population is distributed; secondly, upon the distribution of the density of population over this area; and thirdly, upon the assumed direction of the two bisectors used to locate the median point. The first two of these controlling factors are inherent in the facts but the third depends upon a purely arbitrary decision as to the directions

of the bisectors. Just as the center of area under the bisector definition may be made to travel as indicated in Figure 1 by changing the direction of the bisectors, so the median point of population may be moved by a similar arbitrary decision. After studying carefully a model plate of the form indicated in Figure 1, the writer concluded that an approximate estimate could be made of the direction of a line bisecting the population of the United States which would be at the maximum distance from the median point as located in Bulletin No. 62. The direction was estimated and the position of the corresponding bisector located. This bisector just touches the north-eastern corner of Massachusetts and passes twenty five miles south of Cincinnati. The median point as located in Bulletin No. 62 is seventy miles northward and a little to the westward of Cincinnati. No point of this oblique bisector is within eighty miles of the median point as given in Bulletin No. 62. Another bisector intersecting this one would locate a median point, and it is evident that such a point must be more than eighty miles from the published median point. By arbitrarily changing the direction of the bisectors used to fix it the median point for 1900 can be made therefore to travel about over distances as great as the movement of the properly defined center of population in the last twenty or thirty years.

It may be argued that a knowledge of the exact location of the center of population is of little value, that its location has no appreciable relation of much statistical value to any of the important facts in regard to population. Is this true? In so far as the writer is aware the most important property of the center of population is that its movement from decade to decade indicates the general movement of the population. This is the only important use to which the knowledge of the center has been put in the past, and is absolutely the only use of it made in Bulletin No. 62. Let the relative merits for this purpose of the median point as defined in Bulletin No. 62 and of the center be considered.

As stated in Bulletin No. 62 the movements of the median point "from census to census bear no relation to the movements of the population, since only movements by which bodies of population are transferred across its parallel or its meridian have any influence upon its position." If the whole population of Chicago were suddenly to move to Seattle the median point would remain absolutely stationary. On the other hand, the removal of a hundred men from a certain township in Indiana to the adjoining township on the other side of the median point would move it appreciably. Hence the median point gives no information as to the general movement of the population.

The center of population as fixed by the gravity definition responds to every movement of every individual. It may be shown mathematically, either in the abstract form or from an analysis of the process by which the center is located, that it responds exactly to each movement of every individual by moving in a direction parallel to the movement of the individual, and by an amount exactly proportional to his movement but on a reduced scale fixed by the ratio which this individual bears to the whole population. Thus, if the total population is 75,000,000 and every one stands still except one man the center of population will always move when this one man moves, in a path exactly parallel and similar in every respect to his motion but on the reduced scale of $\frac{1}{75000000}$. If the man moves 1200 miles the center will move about one inch. It is worth while to note in passing that the effect of a given motion on the part of one man upon the position of the center is absolutely independent of the distance of this man from the center. If more than one man moves the movements of the center will be the combined movement due to the different men just as if the movements took place successively instead of simultaneously. Thus if a train moving at thirty miles an hour carries two hundred people the center will respond, provided no other persons move, by moving about five inches each

hour on a path similar in every respect to that followed by the train. If, however, another train were moving at the same time in the reverse direction the movement of the center of population due to this train would partially or wholly neutralize that due to the first train. For, any hour the movement of the center of population is due to the movements of millions of indiviuals, and is the same as if the separate movements of indiviuals had taken place successively and the center had responded on a reduced scale of $\frac{1}{75000000}$ to each in turn. In short the movement of the center of population is an accurate indication on a small scale of the general movement of population.

Or, looking at it from another point of view, the movement of the center of population is an accurate representation of the *average movement of the population.* A movement of a single individual a certain distance is obviously equivalent to a corresponding movement of 75,000,000 people through $\frac{1}{75000000}$ that distance, or the latter motion on the part of the center is the average motion of the whole population.

The movement of the population during a decade is made up not only of the movements of individuals but also of eliminations by death and emigration, and additions by birth and immigration. The effect of the death or emigration of an individual on the position of the center of population defined and computed as indicated in the foregoing is the same as if said individual were first to travel to the center of population and then disappear. Similarly, the effect of a birth or the immigration of an individual is the same as if said individual came into existence at the center of population and then moved to his actual location in the country. This effect of the individuals who appear or disappear from the total population upon the computed position of the center is not so irrational as might seem at first sight. The combined effect of a death in Boston and a birth in San Francisco upon the position of the center should obviously be the

same as if one individual moved from Boston to San Francisco, and this is in turn equivalent to one person going from Boston to the center and disappearing while the other appears at the center and goes to San Francisco. Similarly the combined effect of the landing of an immigrant at New York and the departure of an emigrant from Seattle upon the center of population is and should be the same as if a man moved from Seattle to New York.


I
II
III
IV
V
VI
VII
VIII
IX
X
WASHINGTON
OREGON
CALIFORNIA
NEVADA
IDAHO
MONTANA
WYOMING
UTAH
ARIZONA
COLORADO
NEW MEXICO
NORTH DAKOTA
SOUTH DAKOTA
NEBRASKA
KANSAS
OKLAHOMA
TEXAS
MINNESOTA
IOWA
MISSOURI
ARKANSAS
LOUISIANA
WISCONSIN
ILLINOIS
MICHIGAN
INDIANA
OHIO
KENTUCKY
TENNESSEE
MISSISSIPPI
ALABAMA
GEORGIA
FLORIDA
SOUTH CAROLINA
NORTH CAROLINA
VIRGINIA
WEST VIRGINIA
PENNSYLVANIA
NEW YORK
MAINE
VERMONT
NEW HAMPSHIRE
MASSACHUSETTS
CONNECTICUT
RHODE ISLAND
NEW JERSEY
DELAWARE
MARYLAND
INDIAN TERRITORY